Forty-Sixth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 23-26, 2008

Cayley's Hyperdeterminant, the Principal Minors of a Symmetric Matrix and the Entropy Region of 4 Gaussian Random Variables

Sormeh Shadbakht and Babak Hassibi
Electrical Engineering Department
California Institute of Technology
Pasadena, California 91125
Email: sormeh, hassibi@caltech.edu

***Abstract*— It has recently been shown that there is a connection between Cayley's hypdeterminant and the principal minors of a symmetric matrix. With an eye towards characterizing the entropy region of jointly Gaussian random variables, we obtain three new results on the relationship between Gaussian random variables and the hyperdeterminant. The first is a new (determinant) formula for the $2 \times 2 \times 2$ hyperdeterminant. The second is a new (transparent) proof of the fact that the principal minors of an $n \times n$ symmetric matrix satisfy the $2 \times 2 \times \ldots \times 2$ (n times) hyperdeterminant relations. The third is a minimal set of 5 equations that 15 real numbers must satisfy to be the principal minors of a 4×4 symmetric matrix.**

I. Introduction

Let $X_1, \cdots, X_n$ be n jointly distributed discrete random variables with arbitrary alphabet size N. The vector of all the $2^n - 1$ joint entropies of these random variables is referred to as their "entropy vector" and conversely any $2^n - 1$ dimensional vector whose elements can be regarded as the joint entropies of some n random variables, for some alphabet size N, is called "entropic". The *entropy region* is defined as the region of all possible entropic vectors and is denoted by Γ_n^* [1]. Due to its deep connections with important problems in information theory and probabilistic reasoning such as the capacity of information networks [2][3], or the conditional independence compatibility problem [4], characterizing this region turns out to be of fundamental importance. While it is completely solved for $n = 2, 3$ random variables, the complete characterization for $n \geq 4$ remains an interesting open problem.

The above discussion focused on discrete random variables; however, characterizing the entropy region of a number of continuous random variables is as important. In fact, it has been shown in [5] that there is a correspondence between the continuous and discrete information inequalities and therefore one can characterize one region from the other. Let $\mathcal{N} = \{1, \cdots, n\}$ and for any $\alpha \subseteq \mathcal{N}$, let $H_\alpha = H(X_i, i \in \alpha)$ (or h_α whenever the underlying probability distributions are continuous) be the joint entropies. A valid discrete information inequality of the form $\sum_\alpha a_\alpha H_\alpha \geq 0$ is called *balanced* if for all $i \in \mathcal{N}$, $\sum_{\alpha: i \in \alpha} a_\alpha = 0$. For example $H_1 + H_2 - H_{12} \geq 0$ is balanced and $H_1 \geq 0$ is not.

Theorem 1 (Discrete/continuous information inequalities): [5]

1) A linear continuous information inequality $\sum_\alpha a_\alpha h_\alpha \geq 0$ is valid if and only if its discrete counterpart $\sum_\alpha a_\alpha H_\alpha \geq 0$ is balanced and valid.
2) A linear discrete information inequality $\sum_\alpha a_\alpha H_\alpha \geq 0$ is valid if and only if it can be written as $\sum_\alpha \beta_\alpha H_\alpha + \sum_{i=1}^n r_i (H_{i,i^c} - H_{i^c}) \geq 0$ for some $r_i \geq 0$, where $\sum_\alpha \beta_\alpha h_\alpha \geq 0$ is a valid continuous information inequality (i^c denotes the complement of i in $\mathcal{N}$).

Therefore one can study continuous random variables to determine Γ_n^*. Among all continuous random variables Gaussians are the most natural ones to study first. In fact it turns out that these distributions have interesting properties that make them even more desirable to study.

Let $X_1, \cdots, X_n \in \mathbb{R}^T$ be n jointly distributed zero-mean[1] vector valued Gaussian random variables of dimension T with covariance matrix $R \in \mathbb{R}^{nT \times nT}$. Clearly, R is symmetric, positive semi-definite, and consists of block matrices of size $T \times T$ (corresponding to each random variable). We will allow T to be arbitrary and will therefore consider the *normalized* joint entropy of any subset $\alpha \subseteq \mathcal{N}$ of these random variables

$$\underline{h}_\alpha = \frac{1}{T} \cdot \frac{1}{2} \log \left((2\pi e)^{T|\alpha|} \det R_\alpha \right), \tag{1}$$

where $|\alpha|$ denotes the cardinality of the set α and R_α is the $|\alpha|T \times |\alpha|T$ matrix obtained by keeping those block rows and block columns of R that are indexed by α. Note that our normalization is by the dimensionality of the X_i, i.e., by T, and that we have used $\underline{h}$ to denote normalized entropy. Normalization has the following important consequence.

Theorem 2 (Convexity of the region for $\underline{h}$): The closure of the region of normalized Gaussian entropy vectors is convex [6].

It further turns out that for $n = 2, 3$ random variables, vector-valued Gaussian random variables can be used to obtain the entire entropy region for continuous random variables [6].

This work was supported in part by the National Science Foundation through grant CCF-0729203, by the David and Lucille Packard Foundation, by the Office of Naval Research through a MURI under contract no. N00014-08-1-0747, and by Caltech's Lee Center for Advanced Networking.

[1] Since differential entropy is invariant to shifts there is no point in assuming nonzero means for the X_i.

978-1-4244-2926-4/08/$25.00 ©2008 IEEE

Theorem 3: (Gaussians generate the entropy region for $n = 2, 3$) For two and three random variables, the cone generated by the space of vector-valued Gaussian entropy vectors is the entire entropy region for continuous random variables.

In an effort to characterize the entropy region of discrete random variables, some inner and outer bounds have been established among which the Ingleton bound is the most well-known. The Ingleton inequality was first discovered for the ranks of representable matroids [7]. In fact let $v_1, \cdots, v_n$ be n vector subspaces and $\mathcal{N} = \{1, \cdots, n\}$. Further let $\alpha \subseteq \mathcal{N}$ and r_α be the rank function defined as the dimension of the subspace $\oplus_{i\in\alpha} v_i$. Then for any subsets $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \subseteq \mathcal{N}$, the Ingleton inequality is defined as

$$r_{\alpha_1} + r_{\alpha_2} + r_{\alpha_1\cup\alpha_2\cup\alpha_3} + r_{\alpha_1\cup\alpha_2\cup\alpha_4} + r_{\alpha_3\cup\alpha_4} \\ -r_{\alpha_1\cup\alpha_2} - r_{\alpha_1\cup\alpha_3} - r_{\alpha_1\cup\alpha_4} - r_{\alpha_2\cup\alpha_3} - r_{\alpha_2\cup\alpha_4} \leq 0 \quad (2)$$

Although not all the entropy vectors satisfy this inequality [8], it turns out that certain types of entropy vectors, in particular all the linearly representable (corresponding to linear codes over finite fields) and the abelian group characterizable entropy vectors do and hence fall into this innerbound. An important property of Gaussian random variables is that the entropy vector of 4 jointly Gaussian distributed random variables can be arranged so as to violate the Ingleton bound [6][9].

A. Cayley's Hyperdeterminant

Recall that the entropy of a collection of Gaussian random variables is simply the "log-determinant" of their covariance matrix. Similarly, the entropy of any subset of variables from a collection of Gaussian random variables is simply the "log" of the principal minor of the covariance matrix corresponding to this subset. Therefore one approach to characterizing the entropy region of Gaussians, is to study the determinantal relations of a symmetric positive semi-definite matrix.

For example, consider 3 Gaussian random variables. While the entropy vector of 3 random variables is a 7 dimensional object, there are only 6 free parameters in a symmetric positive semi-definite matrix. Therefore the minors should satisfy a relation. It has very recently been shown that this relation is given by the Cayley's so-called $2 \times 2 \times 2$ "hyperdeterminant" [10]. The hyperdeterminant is a generalization of the determinant concept for matrices to tensors and it was first introduced by Cayley in 1845 [11].

There are a couple of equivalent definitions for the hyperdeterminant among which we choose the definition through the degeneracy of a multilinear form. Consider the following multilinear form of the format $(k_1 + 1) \times (k_2 + 1) \times \ldots \times (k_n + 1)$ in variables $X_1, \ldots, X_n$ where each variable X_j is a vector of length $(k_j + 1)$ with elements in $\mathbb{C}$:

$$f(X_1, X_2, \ldots, X_n) = \\ \sum_{i_1=0}^{k_1}\sum_{i_2=0}^{k_2}\cdots\sum_{i_n=0}^{k_n} a_{i_1,i_2,\ldots,i_n} x_{1,i_1} x_{2,i_2}, \ldots, x_{n,i_n} \quad (3)$$

The multilinear form f is said to be degenerate if and only if there is a non-trivial solution $(X_1, X_2, \ldots, X_n)$ to the following system of partial derivative equations [12]:

$$\frac{\partial f}{\partial x_{j,i}} = 0 \quad \text{for all} \quad j = 1, \ldots, n \quad \text{and} \quad i = 1, \ldots, k_j \quad (4)$$

The unique (up to a scale) irreducible polynomial with integral coefficients in the entries $a_{i_1,i_2,\ldots,i_n}$ of a tensor A that vanishes when f is degenerate is called the hyperdeterminant.

Example (2×2 hyperdeterminant): Consider the 2×2 hyperdeterminant, $f(X_1, X_2) = \sum_{i,j=0}^{1} a_{i,j} x_i y_j$. The multilinear form f is degenerate if there is a non-tirivial solution for X_1, X_2,

$$\frac{\partial f}{\partial x_0} = a_{00}y_0 + a_{01}y_1 = 0 \quad (5)$$
$$\frac{\partial f}{\partial y_0} = a_{00}x_0 + a_{10}x_1 = 0 \quad (6)$$
$$\frac{\partial f}{\partial x_1} = a_{10}y_0 + a_{11}y_1 = 0 \quad (7)$$
$$\frac{\partial f}{\partial y_1} = a_{01}x_0 + a_{11}x_1 = 0 \quad (8)$$

Trying to solve this system of equations, we obtain that,

$$\frac{y_0}{y_1} = \frac{-a_{01}}{a_{00}} = \frac{-a_{11}}{a_{10}} \quad (9)$$
$$\frac{x_0}{x_1} = \frac{-a_{10}}{a_{00}} = \frac{-a_{11}}{a_{01}} \quad (10)$$

We see that a non-trivial solution exists if and only if, $a_{00}a_{11} - a_{10}a_{01} = 0$, i.e. the hyperdeterminant is simply the determinant in this case.

The hyperdeterminant of a $2 \times 2 \times 2$ multilinear form was first computed by Cayley [11] and is as follows:

$$-a_{000}^2a_{111}^2 - a_{100}^2a_{011}^2 - a_{010}^2a_{101}^2 - a_{001}^2a_{110}^2 \\ -4a_{000}a_{110}a_{101}a_{011} - 4a_{100}a_{010}a_{001}a_{111} \\ +2a_{000}a_{100}a_{011}a_{111} + 2a_{000}a_{010}a_{101}a_{111} \\ +2a_{000}a_{001}a_{110}a_{111} + 2a_{100}a_{010}a_{101}a_{011} \\ +2a_{100}a_{001}a_{110}a_{011} + 2a_{010}a_{001}a_{110}a_{101} = 0 \quad (11)$$

In [10] it is further shown that the principal minors of an $n \times n$ symmetric matrix satisfy the $\underbrace{2 \times 2 \times \ldots \times 2}_{n \text{ times}}$ hyperdeterminant. It is thus clear that determining the entropy region of Gaussian random variables is intimately related to Cayley's hyperdeterminant.

It is with this viewpoint in mind that we study the hyperdeterminant in this paper. The paper has three major results. The first is a new determinant formula for the $2\times2\times2$ hyperdeterminant, which is presented in Section II. This may be of interest since computing the hyperdeterminant of higher formats is extremely difficult and our formula may suggest a way of attacking more complicated hyperdeterminants. The second is a novel proof of a main result of [10] that the principal minors of any $n \times n$ symmetric matrix satisfy the $\underbrace{2 \times 2 \times \ldots \times 2}_{n \text{ times}}$ hyperdeterminant. Our proof hinges on identifying a determinant formula for the multilinear form from which the hyperdeterminant arises. This is done in Section III. The third result is a minimal set of five equations that the elements of a real vector in $\mathbb{R}^{15}$ should satisfy to be the principal minors of a 4×4 symmetric matrix. (In contrast, [10] presents a set of 16 equations). This is done in Section IV and gives the relations between the elements of the entropy vector arising from 4 scalar Gaussian random variables.

II. A Formula for the $2 \times 2 \times 2$ Hyperdeterminant

Obtaining an explicit formula for the hyperdeterminant is not an easy task. See, for example, [13] which shows that the $2 \times 2 \times 2 \times 2$ hyperdeterminant consists of 2894276 terms. Here we propose a new formula for (and a method to obtain) the $2 \times 2 \times 2$ hyperdeterminant which might be extendable to hyperdeterminants of larger format.

Theorem 4: (Determinant formula for $2 \times 2 \times 2$ hyperdeterminant) Define

$$B_0 = \begin{bmatrix} a_{000} & a_{100} \\ a_{001} & a_{101} \end{bmatrix}, \ B_1 = \begin{bmatrix} a_{010} & a_{110} \\ a_{011} & a_{111} \end{bmatrix}, \ J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Then the $2 \times 2 \times 2$ hyperdeterminant is given by

$$\det(B_0 J B_1^T - B_1 J B_0^T). \tag{12}$$

Proof: Let f be a multilinear form of the format $2 \times 2 \times 2$,

$$f(X,Y,Z) = \sum_{i,j,k=0}^{1} a_{ijk} x_i y_j z_k \tag{13}$$

Then by the change of variables, $w_0 = x_0 y_0$, $w_1 = x_1 y_0$, $w_2 = x_0 y_1$, $w_3 = x_1 y_1$, the function f can be written as,

$$\begin{aligned} f(X,Y,Z) &= \begin{pmatrix} z_0 & z_1 \end{pmatrix} \begin{pmatrix} a_{000} & a_{100} & a_{010} & a_{110} \\ a_{001} & a_{101} & a_{011} & a_{111} \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\ &\triangleq Z^T \begin{pmatrix} B_0 & B_1 \end{pmatrix} W \end{aligned} \tag{14}$$

To proceed, recall from (4) that the hyperdeterminant of the multilinear form of the format $2 \times 2 \times 2$, vanishes if and only if there is a non-trivial solution (X, Y, Z) to the system of partial derivative equations:

$$\frac{\partial f}{\partial x_i} = 0 \qquad \frac{\partial f}{\partial y_j} = 0 \qquad \frac{\partial f}{\partial z_k} = 0 \qquad i,j,k = 0,1 \tag{15}$$

(a) First we show that if there is a non-trivial solution to the equations (15), then (12) vanishes. By the chain rule $\frac{\partial f}{\partial x_i} = \sum_k \frac{\partial w_k}{\partial x_i} \frac{\partial f}{\partial w_k}$, we can write $\frac{\partial f}{\partial (X,Y)} = \left(\frac{\partial W}{\partial (X,Y)}\right)^T \frac{\partial f}{\partial W}$. Also from (14), $\frac{\partial f}{\partial Z} = \begin{pmatrix} B_0 & B_1 \end{pmatrix} W$. Therefore the degeneracy conditions equivalent with (15) become:

$$\left(\frac{\partial W}{\partial (X,Y)}\right)^T \frac{\partial f}{\partial W} = 0 \tag{16}$$

$$\begin{pmatrix} B_0 & B_1 \end{pmatrix} W = 0 \tag{17}$$

Condition (16) implies that the vector $\frac{\partial f}{\partial W}$ should belong to the null space of $\left(\frac{\partial W}{\partial (X,Y)}\right)^T$.

The following Lemma gives the structure of this null space.

Lemma 1: Null space of the matrix $\left(\frac{\partial W}{\partial (X,Y)}\right)^T$ is characterized by vectors of the form,
$\begin{pmatrix} w_3 & -w_2 & -w_1 & w_0 \end{pmatrix}^T$.

Proof: Let V be a 4×1 vector. Solving for V in the following,

$$\left(\frac{\partial W}{\partial (X,Y)}\right)^T V = \begin{pmatrix} y_0 & 0 & y_1 & 0 \\ 0 & y_0 & 0 & y_1 \\ x_0 & x_1 & 0 & 0 \\ 0 & 0 & x_0 & x_1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = 0 \tag{18}$$

yields the equations:

$$\frac{v_1}{v_3} = \frac{v_2}{v_4} = -\frac{y_1}{y_0} \tag{19}$$

$$\frac{v_1}{v_2} = \frac{v_3}{v_4} = -\frac{x_1}{x_0} \tag{20}$$

Letting $v_4 = x_0 y_0$ characterizes the vectors in the null space up to a scale:

$$\begin{aligned} V^T &= \begin{pmatrix} x_1 y_1 & -x_0 y_1 & -x_1 y_0 & x_0 y_0 \end{pmatrix}^T \\ &= \begin{pmatrix} w_3 & -w_2 & -w_1 & w_0 \end{pmatrix}^T \end{aligned} \tag{21}$$

■

Going back to the proof of Theorem 4, using Lemma 1 we conclude that we should have, $\frac{\partial f}{\partial Z} = \begin{pmatrix} B_0 & B_1 \end{pmatrix} W = 0$ and for an arbitrary non-zero scalar α, $\frac{\partial f}{\partial W} = \begin{pmatrix} B_0 & B_1 \end{pmatrix}^T Z = \alpha \begin{pmatrix} w_3 & -w_2 & -w_1 & w_0 \end{pmatrix}^T$. Putting these two equations into matrix form we can further write the following:

$$\begin{pmatrix} 0 & 0 & {B_0}^T \\ 0 & 0 & {B_1}^T \\ B_0 & B_1 & 0 \end{pmatrix} \begin{pmatrix} W \\ Z \end{pmatrix} = \alpha \begin{pmatrix} w_3 \\ -w_2 \\ -w_1 \\ w_0 \\ 0 \\ 0 \end{pmatrix} \tag{22}$$

or in other form:

$$\begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \alpha \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & {B_0}^T \\ \alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & {B_1}^T \\ B_0 & B_1 & 0 \end{pmatrix} \begin{pmatrix} W \\ Z \end{pmatrix} = 0 \tag{23}$$

A non-trivial solution for X, Y, Z and hence for W, Z requires:

$$\begin{aligned} &\det\left(\begin{pmatrix} B_0 & B_1 \end{pmatrix} \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix} \begin{pmatrix} {B_0}^T \\ {B_1}^T \end{pmatrix} \right) \\ &\qquad = \det(B_0 J {B_1}^T - B_1 J {B_0}^T) = 0 \end{aligned} \tag{24}$$

Note that the explicit calculation of (24) gives the hyperdeterminant formula of the form $2 \times 2 \times 2$ stated in equation (11) as expected.

(b) Conversely suppose that (24) vanishes and therefore there is a non-trivial solution for W and Z in (23). To prove that there is also a non-trivial solution to (15), we need to show that such X, Y and Z exist so that (16) and (17) hold. It is not hard to see that this is only possible if $W = \begin{pmatrix} w_0 & w_1 & w_2 & w_3 \end{pmatrix}^T$ in (23) has the property,

$$\frac{w_0}{w_2} = \frac{w_1}{w_3} \tag{25}$$

In the following we show that the solution of (23) in fact satisfies relation (25). Let $p = \begin{pmatrix} w_0 & w_1 \end{pmatrix}^T$ and $q = \begin{pmatrix} w_2 & w_3 \end{pmatrix}^T$. Then from (23) we obtain:

$$\alpha Jq \quad + \quad {B_0}^T Z = 0 \tag{26}$$
$$-\alpha Jp \quad + \quad {B_1}^T Z = 0 \tag{27}$$
$$B_0 p \quad + \quad B_1 q = 0 \tag{28}$$

Multiplying the first equation by p^T and the second one by q^T and adding them together we obtain,

$$\alpha(p^T Jq - q^T Jp) + (p^T {B_0}^T + q^T {B_1}^T)Z = 0 \tag{29}$$

which by the use of (28) simplifies to:

$$p^T Jq = q^T Jp \tag{30}$$

Noting that $p^T Jq = (p^T Jq)^T = -q^T Jp$ gives,

$$p^T Jq = q^T Jp = 0 \tag{31}$$

(25) then follows immediately from (31) by substituting for p and q. ■

III. Minors of a Symmetric Matrix Satisfy the Hyperdeterminant

It has recently been shown in [10] that the principal minors of a symmetric matrix satisfy the hyperdeterminant relations. There this was found by either checking or explicitly computing the determinant of a 3×3 matrix in terms of the other minors and noticing that it satisfied the $2 \times 2 \times 2$ hyperdeterminant. In this section we give an explanation of why this relation holds for the principal minors of a symmetric matrix. The key ingredient is by identifying a simple determinant formula for the multilinear form (3) when the coefficients $a_{i_1,i_2,\ldots,i_n}$ are the minors of an $n \times n$ symmetric matrix.

Lemma 2: Let the elements of the tensor $A = [a_{i_1,i_2,\ldots,i_n}]$ be the principal minors of a symmetric $n \times n$ matrix $\tilde{A}$, then the following multilinear form of the format $2 \times 2 \times \ldots \times 2$ (n times),

$$f(X_1, X_2, \ldots, X_n) = \sum_{i_1,i_2,\ldots,i_n=0}^{1} a_{i_1,i_2,\ldots,i_n} x_{1,i_1} x_{2,i_2}, \ldots, x_{n,i_n} \tag{32}$$

can be rewritten as the determinant of the matrix M, i.e. $f(X_1, X_2, \ldots, X_n) = \det(M)$ where M is the following matrix:

$$M = \begin{pmatrix} x_{1,0} & 0 & \ldots & 0 \\ 0 & x_{2,0} & \ldots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & x_{n,0} \end{pmatrix} + \begin{pmatrix} x_{1,1} & 0 & \ldots & 0 \\ 0 & x_{2,1} & \ldots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & x_{n,1} \end{pmatrix} \tilde{A}$$
$$\triangleq N_1 + N_2 \tilde{A} \tag{33}$$

Proof: Let $(p_1 \ldots p_n)$ be a realization of $\{0,1\}^n$. For $j = 1, \ldots, n$ in (33), let the variables $x_{j,p_j} = 1$ and the rest of the variables be zero. Then it can be easily seen that $f(X_1, X_2, \ldots, X_k)$ in (33) will be equal to the minor of the matrix $\tilde{A}$ obtained by choosing the set of rows and columns $\alpha \subseteq \{1, \ldots, n\}$ such that $p_j = 1$ for all $j \in \alpha$. This is nothing but the coefficient $a_{p_1,p_2,\ldots,p_n}$ in (32). ■

Lemma 3 (Partial derivatives of $\det M$*):* Let $\alpha = \{1, \ldots, n\} \setminus j$. Computing the partial derivatives of the $\det M$ gives ,

$$\frac{\partial \det M}{\partial x_{j,0}} = \det M_{\alpha,\alpha} \tag{34}$$
$$\frac{\partial \det M}{\partial x_{j,1}} = \det \tilde{A} \det M'_{\alpha,\alpha} \tag{35}$$
$$\textit{where} \quad M' = N_1 \tilde{A}^{-1} + N_2$$

Proof: (34) can be proved by direct calculation. For (35) note that:

$$\begin{aligned} \frac{\partial \det M}{\partial x_{j,1}} &= \frac{\partial}{\partial x_{j,1}} \det[(N_1 \tilde{A}^{-1} + N_2)\tilde{A}] \\ &= \det \tilde{A} \frac{\partial \det M'}{\partial x_{j,1}} \end{aligned} \tag{36}$$

Using (34), and the above equation, (35) follows immediately. ■

Now we obtain the necessary and sufficient condition for the minors of $\tilde{A}$ to satisfy the hyperdeterminant in terms of the rank of the matrix M:

Lemma 4 (rank of M*):* The minors of the matrix $\tilde{A}$ satisfy the hyperdeterminant equation if and only if rank of M in (33) is at most $n-2$.

Proof: To satisfy the hyperdeterminant, we require (34) and (35) to be equal to zero simultaneously. However vanishing of (34) implies that all the $(n-1)\times(n-1)$ principal minors of M are zero and since M is symmetric this means that M should be of rank at most $n-2$. Moreover assuming $\tilde{A}$ to be nonsingular, (35) being zero requires M' to be low rank as well. Noting that $M' = M\tilde{A}^{-1}$, (35) being zero follows directly from vanishing of (34). Conversely suppose that M has rank of at most $n-2$. Then both (34) and (35) vanish and the multilinear form (32) becomes degenerate which means the coefficients $a_{i_1,i_2,\ldots,i_n}$ i.e. the principal minors of the matrix, will satisfy the hyperdeterminant. ■

Theorem 5 (hyperdeterminant and the principal minors): The principal minors of an $n \times n$ symmetric matrix $\tilde{A}$

satisfy the hyperdeterminants of the format $2 \times 2 \ldots \times 2$ (k times) for all $k \leq n$.

Proof: Clearly it suffices to show that the minors satisfy the $2 \times 2 \ldots \times 2$ (n times) hyperdeterminant. Recall that for the tensor of coefficients $a_{i_1,i_2,\ldots,i_n}$ in the multilinear form (3) to satisfy the hyperdeterminant relation, there must exist a non-trivial solution to make all the partial derivatives of f with respect to its variables zero. Therefore using Lemma (4), we need to find a non-trivial solution to make the matrix M of rank at most $n-2$. In the following we will show that such a solution always exists.

First we find a non-trivial solution in the case of 3 variables and then extend it to the the case where there are n variables. For 3 variables, the matrix M should be rank 1 or equivalently all the columns be multiples of one another. Enforcing this condition results in 3 equations for 6 unknowns. Therefore without loss of generality we let $x_{j,1} = 1$. Making the columns of M proportional, gives:

$$\frac{x_{1,0} + a_{11}}{a_{12}} = \frac{a_{12}}{x_{2,0} + a_{22}} = \frac{a_{13}}{a_{23}} \tag{37}$$

$$\frac{x_{3,0} + a_{33}}{a_{23}} = \frac{a_{13}}{a_{12}} \tag{38}$$

If $\overline{x}_i = (x_{i,0}, x_{i,1})$, then the solution to the above equations is clearly as follows:

$$\begin{aligned}\overline{x}_1 &= (\frac{a_{12}a_{13} - a_{11}a_{23}}{a_{23}}, 1)\\ \overline{x}_2 &= (\frac{a_{23}a_{12} - a_{13}a_{22}}{a_{13}}, 1)\\ \overline{x}_3 &= (\frac{a_{13}a_{23} - a_{12}a_{33}}{a_{12}}, 1)\end{aligned} \tag{39}$$

Now for the general case of n variables, let $\overline{x}_1, \overline{x}_2, \overline{x}_3$ be as (39) and for $j > 3$, $\overline{x}_j = (1, 0)$. It can be easily checked that this solution makes the matrix M of rank $n-2$. ■

IV. Minimal conditions for the elements of a 15−dimensional vector to be the principal minors of a symmetric 4×4 matrix

In this section we shall give a minimal set of equations that the elements of a 15-dimensional vector must satisfy to be the principal minors of a 4×4 symmetric matrix. These can be used as the starting point to determining the entropy region of 4 jointly Gaussian scalar random variables. (By contrast [10] gives a nonminimal set of 16 equations).

Let the elements of the vector $A \in \mathbb{R}^{2^n-1}$ be denoted by A_α, $\alpha \subseteq \{1, \ldots, n\}$. An interesting problem is to find the minimal set of conditions under which the vector A can be considered as the vector of all principal minors of a symmetric $n \times n$ matrix. This problem has been addressed before in [10], [14]. Here we propose the minimal set of such conditions for $n = 4$.

Roughly speaking there are 15 variables in the vector A and only 10 parameters in a symmetric 4×4 matrix. Therefore if the elements of A can be considered as the minors of a 4×4 symmetric matrix, one suspects that there should be 5 constraints on the elements of A. In fact we find 5 such constraints and to prove their sufficiency, we show that for a given vector A and under such constraints one can construct the symmetric matrix $\tilde{A} = [a_{ij}]$. Let

$$g_{ijk} = A_{ijk} - A_iA_{jk} - A_jA_{ik} - A_kA_{ij} + 2A_iA_jA_k \tag{40}$$

Theorem 6: The minimal set of necessary and sufficient conditions for the elements of the vector A to be the principal minors of a symmetric 4×4 matrix consists of three hyperdeterminant equations, one consistency of the signs of g_{ijk} and the determinant identity of the 4×4 matrix:

$$g_{123}^2 = 4(A_1A_2 - A_{12})(A_2A_3 - A_{23})(A_1A_3 - A_{13}) \tag{41}$$

$$g_{124}^2 = 4(A_1A_2 - A_{12})(A_2A_4 - A_{24})(A_1A_4 - A_{14}) \tag{42}$$

$$g_{134}^2 = 4(A_1A_3 - A_{13})(A_3A_4 - A_{34})(A_1A_4 - A_{14}) \tag{43}$$

$$g_{123}g_{124}g_{134} = 4(A_1A_2 - A_{12})(A_1A_3 - A_{13})(A_1A_4 - A_{14})g_{234} \tag{44}$$

$$\begin{aligned}A_{1234} = &-\frac{1}{2}\sum_{\substack{i,j\in\{1,2,3\}\\ k,l\in\{1,2,3,4\}\setminus\{i,j\}}} \frac{g_{ijk}g_{ijl}}{A_iA_j - A_{ij}} + A_1g_{234} + A_2g_{134}\\ &+A_3g_{124} + A_4g_{123} - 2A_1A_2A_3A_4 + A_{12}A_{34}\\ &+A_{13}A_{24} + A_{14}A_{23}\end{aligned} \tag{45}$$

Proof: Necessity is straightforward to show. For sufficiency, first note that all the elements of $\tilde{A}$ can be determined up to a sign from the A_i and A_{ij} elements of the vector A.

$$a_{ii} = A_i \tag{46}$$

$$a_{ij}^2 = a_{ii}a_{jj} - A_{ij} = A_iA_j - A_{ij} \tag{47}$$

It remains to choose the signs of all the off-diagonals in such a way that the 3×3 and 4×4 minors of $\tilde{A}$ will correspond to A_{ijk} and A_{1234}. First let's consider the 3×3 minors. Assuming $\tilde{A}$ to be the symmetric matrix with minors corresponding to elements of A, a direct calculation of a 3×3 principal minor with rows and columns indexed by $\{i, j, k\}$, gives:

$$\begin{aligned}A_{ijk} &= a_{ii}a_{jj}a_{kk} - a_{ii}a_{jk}^2 - a_{jj}a_{ik}^2 - a_{kk}a_{ij}^2 + 2a_{ij}a_{jk}a_{ik}\\ &= -2A_iA_jA_k + A_iA_{jk} + A_jA_{ik} + A_kA_{ij}\\ &\quad \pm 2\sqrt{(A_iA_j - A_{ij})(A_iA_k - A_{ik})(A_jA_k - A_{jk})}\end{aligned} \tag{48}$$

which can be written as:

$$\begin{aligned}g_{ijk} &= 2a_{ij}a_{jk}a_{ik}\\ &= \pm 2\sqrt{(A_iA_j - A_{ij})(A_iA_k - A_{ik})(A_jA_k - A_{jk})}\end{aligned} \tag{49}$$

Note that although the sign ambiguities of the 3 off-diagonal elements in a 3×3 minor imply 8 possible matrices, the determinant of a 3×3 matrix depends only on the sign of the product of the off-diagonal terms or in other words the parity of g_{ijk}.

Squaring both sides yields the hyperdeterminant relation (11) [2]. There are four such hyperdeterminants for a 4×4 matrix

[2]To see the equivalence with (11), consider one of the hyperdeterminants, e.g. (41) and let the elements of the tensor in the multilinear form (3) be the principal minors A_α such that a_{ijk} is mapped to A_α where $\alpha = \{1 \times i, 2 \times j, 3 \times k\} \setminus \{0\}$.

each corresponding to a 3×3 minor,

$$g_{ijk}^2 = 4{a_{ij}}^2{a_{ik}}^2{a_{jk}}^2 \qquad i,j,k \in \{1,2,3,4\} \tag{50}$$

(50) for all permutations of $\{i,j,k\}$ assures that there is a parity choice for the four g_{ijk} such that all the A_{ijk} will correspond to the 3×3 minors of $\tilde{A}$. However what we require next is the consistency of the parities. In other words there should exist at least one sign assignment of the off-diagonal terms that results in the assumed parities of g_{ijk}. Multiplication of the three of g_{ijk} gives:

$$g_{ijk}g_{ijl}g_{ikl} = 4a_{ij}^2a_{ik}^2a_{il}^2g_{jkl} \tag{51}$$

In other words, once the parities of the three out of four g_{ijk} are determined the last one should be consistent with them through (51). Considering one of these equations, i.e. a particular permutation of $\{i,j,k\}$ is sufficient for our purpose,

$$g_{123}g_{124}g_{134} = 4a_{12}^2a_{13}^2a_{14}^2g_{234} \tag{52}$$

It only remains to insist that the whole determinant of the constructed matrix be equal to A_{1234}. This is guaranteed through (45) which is obtained by direct calculation of the 4×4 determinant. Noting that, one hyperdeterminant equation, for example,

$$g_{234}^2 = 4(A_2A_3 - A_{23})(A_3A_4 - A_{34})(A_2A_4 - A_{24}) \tag{53}$$

can be obtained from the other three hyperdeterminants, i.e. (41),(42) and (43) and the parity consistency condition (52), leaves 5 equations of (41) to (45) through which we can construct the matrix $\tilde{A}$. ■

V. Conclusion

Studying the principal minor relations are crucial to characterizing the entropy region of Gaussian random variables as an interesting subclass of continuous random variables. With this viewpoint, these relations especially the hyperdeterminant were studied in this paper. In particular by giving a determinant formula for a multilinear form, we gave a transparent proof that the hyperdeterminant relation is satisfied by the principal minors of an $n \times n$ symmetric matrix. Moreover we obtained a closed form for the $2 \times 2 \times 2$ hyperdeterminant which might be extendible to higher order formats and is an interesting problem even on its own. Finally a minimal set of 5 necessary and sufficient conditions for 15 numbers to be the principal minor of a symmetric matrix were presented.

References

[1] R.W. Yeung, "A framework for linear information inequalities," *IEEE Trans. on Information Theory*, vol. 43, no. 6, pp. 1924–1934, 1997.

[2] B. Hassibi and S. Shadbakht, "Normalized entropy vectors, network information theory and convex optimization," in *Information theory workshop, Bergen, Norway*, 2007.

[3] X. Yan, R.W. Yeung, and Z. Zhang, "The capacity region for multi-source multi-sink network coding," in *IEEE Int. Symp. on Inf. Theory (ISIT)*, 2007, pp. 116–120.

[4] F. Matus and M. Studeny, "Conditional independences among four random variables I," *Combin., Prob. Comput.*, vol. 4, pp. 269–278, 1995.

[5] T. H. Chan, "Balanced information inequalities," *IEEE Trans. on Information Theory*, vol. 49, no. 12, pp. 3261–3267, 2003.

[6] B. Hassibi and S. Shadbakht, "The entropy region for three gaussian random variables," in *IEEE Int. Symp. on Inf. Theory (ISIT)*, 2008.

[7] A. W. Ingleton, "Representation of matroids," in *Combinatorial mathematics and its applications, D. Welsh, Ed. London: Academic Press*, 1971, pp. 149–167.

[8] Z. Zhang and R. Yeung, "On characterization of entropy function via information inequalities," *IEEE Trans. on Information Theory*, vol. 44, no. 4, pp. 1440–1452, 1998.

[9] Charles R. Johnson and Wayne W. Barrett, "Determinantal inequalities for positive definite matrices," *Discrete Mathematics*, vol. 119, pp. 97–106, 1993.

[10] O. Holtz and B. Sturmfels, "Hyperdeterminantal relations among symmetric principal minors," *Journal of Algebra*, vol. 316, pp. 634–648, 2007.

[11] A. Cayley, "On the theory of linear transformations," in *Cambridge Math. J. 4*, 1845, pp. 1–16.

[12] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants and multidimensional determinants*, Mathematics: Theory and Applications., 1994.

[13] Peter Huggins, Bernd Sturmfels, Josephine Yu, and Debbie S. Yuster, "The hyperdeterminant and triangulations of the 4-cube," *Mathematics of Computation*, vol. 77, no. 263, pp. 1653–1679, 2008.

[14] K. Griffin and Tsatsomeros M. J, "Principal minors, part II: The principal minor assignment problem," *Linear Algebra and its applications*, vol. 419, pp. 125–171, 2006.